American Journal of Bioinformatics Research 2014, 4(2): 23-32
DOI: 10.5923/j.bioinformatics.20140402.01

A Comprehensive *in silico* DNA Sequence Analysis of Sperm Surface ADAM Genes Collected from RefSeq Database

Ayman Sabry[1,2,*], Alaa Ahmed Mohamed[1,2,3], Nabil Said Awad[1]

[1]Biotechnology and Genetic Engineering Unit, Scientific Research Deanship, Taif University, TAIF, Zip Code 21974, Post Box 888 KSA
[2]Cell Biology Department, National Research Cen- ter, Dokki, Giza, Egypt
[3]Animal Reproduction and AI, Veterinary Research Division, National Research Center, Dokki, Giza, Egypt

Abstract The objective of the present study is to provide a comprehensive *in silco* sequence analyses of sperm-surface ADAM genes using the curated and nonredundant RefSeq database. 36 complete refseq CDS (coding sequence) of 9 members of ADAM gene family namely ADAM1,2,3,4,5,6,18,24 and 32 were obtained to investigate its evolution and differentiation within and among species. Among the 9 sperm-surface genes ADAM1 has the longest CDS length, the highest GC% and largest variation in base composition. Measurement of polymorphism and genetic diversity (e.g. Number of haplotypes, nucleotide diversity (π) and average number of nucleotide diversity) varied greatly among the sperm-surface ADAM genes. These large variations are evidence for the effect of selection pressure on these genes. The phylogenetic analysis displayed clearly the resolved relationship of sperm-surface ADAM genes and most of bootstrap support were high. Sperm-surface ADAM genes were classified into 6 clades where, ADAM1,2,3,4,5,6,18,24 and ADAM32 of *B.taurus* and *H. sapiens* categorizing to one large lineage where ADAM32 of *R. norvegicus* and *M. musculus* belonging to another separate lineage. In conclusion, the *in silico* analysis of the 9 sperm-surface ADAM genes showed a great deal of variation among and within this genes indicating the presence of localized signals of selection pressure on these genes. Moreover, ADAM1,4,6 and 24 have no human orthologues. More importantly, our results suggested that ADAM1,2,3,4,5,6, 18,24 and ADAM32 of *B.taurus* and H. sapiens were descendent form one ancient ancestor, where ADAM32 of *R. norvegicus* and *M. musculus* have another ancestor.

Keywords Sperm-surface, ADAM genes, in silco Sequence analysis, DNA sequence diversity, RefSeq

1. Introduction

The interaction of mammalian spermatozoon with the oocytee's extracellular matrix or zona pellucid is critical first step towards successful fertilization. Important key players in this extracellular interaction are ADAM (A Disintegrin And Metalloproteinase) genes. The ADAM gene family comprises 35 characterized genes in mammals, with about 18 genes are known to be expressed exclusively or predominantly in the male reproductive tissue (Wolfsberg *et al.*, 1995b, Blobel, 1997, Black & White, 1998, Primakoff & Myles, 2000, Seals & Courtneidge, 2003, Cho, 2005 and Grayson & Civetta, 2013). ADAM gene family was unearthed during the study of sperm and egg merger that start off zygote development (Blobel *et al.*, 1992, Wolfsberg *et al.*, 1993 and Wolfsberg *et al.*, 1995a). ADAMs were found to have multiple and diverse functions, both tissue-specific as well as ubiquitous patterns of expression and common evolutionary history (Finn & Civetta, 2010). The analysis of ADAM family evolution among mammals has found faster divergence of genes expressed in testes (Civetta, 2003, Glassey & Civetta, 2004, Finn & Civetta, 2010 and Morgan et al., 2010).

In mammals ADAM1,2,3,4,5,6,18,24 and 32 are best characterized in terms of their role during fertilization. The first three genes play a role during sperm migration and zona pellucida (ZP) binding as well as egg membrane recognition and fusion (Blobel *et al.*, 1992, Wolfsberg *et al.*, 1995a, Primakoff & Myles, 2000 and Evans, 2002). For example, knockout mice for ADAM2 and ADAM3 show drastic decreases in sperm aggregation, a trait that has been suggested to confer sperm with competitive advantages (Moore *et al.*, 2002, Fisher & Hoekstra, 2010 and Han *et al.*, 2010). ADAM3 knockouts males were infertile due to deficiencies in sperm-ZP interactions, and more importantly, sperm migration into the oviduct (Shamsadin *et al.*, 1999, Nishimura *et al.*, 2001 and Yamaguchi *et al.*, 2009).

* Corresponding author:
amsabry@gmail.com (Ayman Sabry)
Published online at http://journal.sapub.org/bioinformatics

ADAM2 knock- outs also significantly affect reproductive success. In vivo, ADAM2 null mice have a fertility rate 50 times lower than the wild-type. This drop in fertility once again does not appear to be the result of a single process, but is instead a combination of deficiencies in sperm-egg fusion, sperm-egg binding, spermZP binding and sperm migration (Cho *et al.*, 1998). ADAM1a knockouts result in sperm unable to migrate to the egg; in vivo the knockout produces an infertile phenotype but in vitro, sperm are able to fertilize eggs. ADAM1b knockouts appear to produce normal sperm but affect the levels of ADAM2 on mature sperm (Nishimura *et al.*, 2004 and Kim *et al.*, 2006).

ADAM4,5,6,18,24, and 32 genes are also sperm surface genes while other ADAMs have been identified only as testis expressed (Frayne *et al.*, 2002, Kim *et al.*, 2005 and Zhu *et al.*, 2009). Six of the sperm surface genes (Adams 1 to 6) assemble into functional complexes. Currently, there is evidence for three sperm-specific complexes (ADAM2-ADAM3- ADAM4, ADAM2-ADAM3-ADAM5, and ADAM2-ADAM3- ADAM6), two testis-specific complexes (ADAM1a- ADAM2, and ADAM2-ADAM3), and one complex common to both (ADAM1b-ADAM2) (Cho, 2012). All complexes require at least ADAM2 and/or ADAM3, if not both, and their interactions appear to be central for a variety of sperm functional adaptations to fertility in mice.

From molecular evolutionary stand point, most of protein coding genes have been found to evolve under purifying selection, but genes that function in perception, immunity and reproduction are often fast-evolving exceptions to this rule (Voight *et al.*, 2006, Kosiol *et al.*, 2008 Koonin & Wolf, 2010. Reproductive genes, such as those that code for species-specific fertilization proteins, male accessory gland proteins, and sperm proteins have been shown to exhibit rapid evolution in taxa as diverse as invertebrates, mammals and plants (Swanson & Vacquier, 2002, A, 2003, Clark et al. , 2006, Panhuis et al., 2006, Turner & Hoekstra, 2008 and Dorus *et al.*, 2010).

One of the distinguished NCBI projects is Reference Sequence (RefSeq) database (http://www.ncbi.nlm.nih.gov/RefSeq/). RefSeq is a public database of nucleotide and protein sequences with corresponding features and bibliographic annotations. The RefSeq database is built and distributed by the NCBI. NCBI builds RefSeq from the sequence data available in the archival database GenBank (Benson *et al.*, 2005). The RefSeq collection is unique in providing a curated, nonredundant, explicitly linked nucleotide and protein database representing significant taxonomic diversity. The RefSeq collection is derived from the primary submissions available in GenBank. GenBank is a redundant archival database that represents sequence information generated at different times, and may represent several alternate views of the protein, names or other information. In contrast, RefSeq represents a nearly non-redundant collection that is a synthesis and summary of available information, and represents the current view of the sequence information, names and other annotations (Kim *et al.*, 2005).

The objective of the present study is to provide a comprehensive phylogenetic and sequence analyses of sperm-surface ADAM genes using the curated and non-redundant RefSeq database. 36 complete refseq CDS (coding sequence) of 9 members of ADAM gene family namely ADAM1,2,3,4,5,6,18,24 and 32 were obtained to investigate its evolution and differentiation within and among species.

2. Materials and Methods

Data Collection

Table 1. Accession numbers, sequence length, percent of GC-content and stop codons of 36 CDS for 9 sperm-surface ADAM genes

Gene	Species	Accession No.	Sequence Length bp	GC cont.%	Stop codon
Adam1	M. mulatta	NM 001195734.1	2736	51	TAG
	M. musculus	NM 172126.2(a)	2376	53	TGA
	M. musculus	NM 172125.2(b)	2421	52	TAA
	R. norvegicus	NM 020078.1	2370	55	TGA
	B. taurus	NM 001206471.1	2454	50	TAA
Adam2	H. sapiens	NM 001464.4 (V1)	2208	39	TAG
	H. sapiens	NM 001278113.1(V2)	2151	39	TAG
	H. sapiens	NM 001278114.1(V3)	2019	39	TAG
	M. musculus	NM 009618.2	2208	46	TAG
	B. taurus	NM 174228.1	2238	40	TAG
	S. scrofa	NM 213957.1	2208	39	TAA
	M. fascicularis	NM 001283853.1	2208	39	TAG
	O. cuniculus	NM 001082677.1	2256	41	TAG
	R. norvegicus	NM 020077.1	2220	47	TAG
	C. porcellus	NM 001172910.1	2208	44	TAG
Adam3	M. musculus	NM 009619.4	2469	45	TAA
	R. norvegicus	NM 020302.1	2223	45	TAG
Adam4	M. musculus	NM 009620.1	2292	43	TGA
	R. norvegicus	NM 020305.1	2268	43	TAA
Adam5	M. musculus	NM 001272058.1(V1)	2316	42	TAA
	M. musculus	NM 007401.3 (V2)	2100	42	TGA
	M. musculus	NM 007401.3(V3)	2256	42	TAA
	M. musculus	NM 001272059.1(V4)	2088	43	TGA
	M. fascicularis	NM 001283728.1	2271	38	TAA
	R. norvegicus	NM 020303.1	2130	44	TAA
	C. porcellus	NM 001173099.1	2334	40	TAA
Adam6	M. musculus	NM 174885.3	2265	44	TAA
	O. cuniculus	NM 001165916.1	2196	47	TAA
	R. norvegicus	NM 138906.1	2256	45	TAA
Adam18	H. sapiens	NM 001190956.1	2220	37	TAA
	M. musculus	NM 010084.2	2160	45	TGA
Adam24	M. musculus	NM 010086.4	2286	42	TAG
Adam32	H. sapiens	NM 145004.5	2364	39	TAG
	B. taurus	NM 001046250.1	2244	40	TAA
	M. musculus	NM 153397.2	2265	43	TAA
	R. norvegicus	NM 001170582.1	2259	45	TAA

In the present study data were obtained from the NCBI Reference Sequence (RefSeq) database (http://www.ncbi.nlm.nih.gov/refseq). RefSeq provides a curated non-redundant sequence database of genomes, transcripts and proteins (Pruitt et al., 2005). The search for ADAM genes sequence were restricted to 9 sperm-surface ADAM genes namely, Adam 1,2,3,4,5,6,18,24 and 32 genes. Only complete, verified, experimentally proofed and non-predictable CDS were considered. The data collection resulted in 36 sequences for the 9 selected genes. Detailed of the 36 genes, species and accession numbers are presented in table (1).

In silico sequence analysis

Nucleotides sequences were analysed and translation of nucleotides into amino acid sequence were carried out using Biostrings package (Pages *et al.*, 2013) under the R Project for Statistical Computing (R Core Team, 2013) DnaSP (version 5.10.01) software was used to analyze the haplotype diversity (H_d), the average number of nucleotide differences, the average number of nucleotide differences (Tajima, 1983), the nucleotide diversity (π). The polymorphic site (S), the singleton variable sites (SP), and the parsimony informative sites (PIP) for each gene, and the average number of nucleotide substitutions per site between species (D_{xy}) (Lynch & Crease, 1990) The phylogenetic analysis was carried out using Neighbour jount method (Saitou & Nei, 1987) implemented in ape 3.0 package (Paradis *et al.*, 2004).

3. Results

Sequence Variations & GC-content

36 CDS sequences of ADAM1,2,3,4,5,6,18,24 and 32 genes were obtained from NCBI Reference Sequence (RefSeq) database (http://www.ncbi.nlm.nih.gov/refseq). A detailed list of the NCBI refseq accession numbers for the 36 sperm-surface Adam genes as well as percentage GC-content and length of CDS are presented in table (1). The CDS length varied substantially within each of ADAM genes, even for species with close taxonomic relationship (e.g. *M. musculus* and *N. norvegicus*) as well as among the 9 genes. The longest sequence was observed in ADAM1, for the 4 species CDS length ranged from 2370 to 2736 bp, where ADAM2 has shortest, and its length ranged form 2019 to 2238 bp. Among the 8 species of ADAM2, *homo sapiens* had the shortest length for its 3 variants, that is 2019, 2151 and 2208 bp. Figure (1) shows the base composition of the 9 ADAM genes, with exception of ADAM1 no noticeable differences are observed. ADAM1 has the lowest percentage of A and T bases but also ADAM1 has the highest percentages of G and C. These differences are more pronounced when GC-content was considered. The GC-content of all studied sequences ranged from 37 to 55%. Figure (2) shows that the range of variation within each of the 9 genes differs from one gene to another. The range of GC-content% for ADAM 1 was the highest (50–55%) where ADAM2 was the lowest (39–47%). However, ADAM3 and ADAM4 showed exactly similar GC-content, although genes have different CDS length for both of *M. musculus* and *N. norvegicus*.

Variation of stop Codon

The three types of stop codons were observed in ADAM1, even for the two variants of *M. musculu*s two different stop codon (TGA & TAA) were also observed (table 1). For ADAM2 most species have TAG as stop codon, where only *S.scrofa* has TAA stop codon. For ADAM3,4,5,6,18 and 32, most species have TAA stop codon, where the presence of TAG or TGA is few species could be indicate a presence of mutation in the stop codon.

ADAM Gene Phylogeny

Phylogenetic analysis of the 9 sperm-surface ADAM genes was carried out to reconstruct phylogenetic tree using Neighbor-Joining (NJ) method (Figure 3). This phylogenetic tree displayed clearly the resolved relationship of sperm-surface ADAM genes and most of bootstrap support were high.

Sperm-surface ADAM genes were classified into 6 clades where, ADAM1,2,3,4,5,6,18,24 and ADAM32 of *B.taurus* and H. sapiens categorizing to one large lineage where ADAM32 of *R. norvegicus* and *M. musculus* belonging to another separate lineage. This result suggests that ADAM1,2,3,4,5,6, 18,24 and ADAM32 of *B.taurus* and H. sapiens were descendent form one ancient ancestor, where ADAM32 of *R. norvegicus* and *M. musculus* have another ancestor.

ADAM2 genes of all studied species were clustered into a monophyletic group, which suggesting that ADAM2 is the most conserved gene across species. ADAM1,4,6 and 24 constituted another monophyletic group. ADAM3,5 and 18 were comprised in another separate clade.

Polymorphism and Genetic Diversity among species

Only one sequence for ADAM24 gene was obtained form RefSeq database, therefore, this sequence was ruled out from this analysis. Sequences of the rest 8 members of ADAM gene family were aligned using both Clustalx and Maft build in ape package (Paradis *et al.,* 2004). As the alignment results of the two software were found to be identical, the robustness of the alignment method is ensured. For total number of parsimony-informative sites (i.e. sites that have a minimum of two nucleotides that are present at least twice), each of ADAM3, 4,6,18 showed no parsimony-informative sites, that is, ADAM32 had the smallest number of parsimony-informative sites where ADAM2 had the largest number of informative sites (709). All polymorphic sites for ADAM3,4,6, and 18 were found to be noninformative (singleton), for the rest of ADAM genes ADAM2 had the lowest number of singleton variable sites (376) where ADAM1 had the largest number (645).

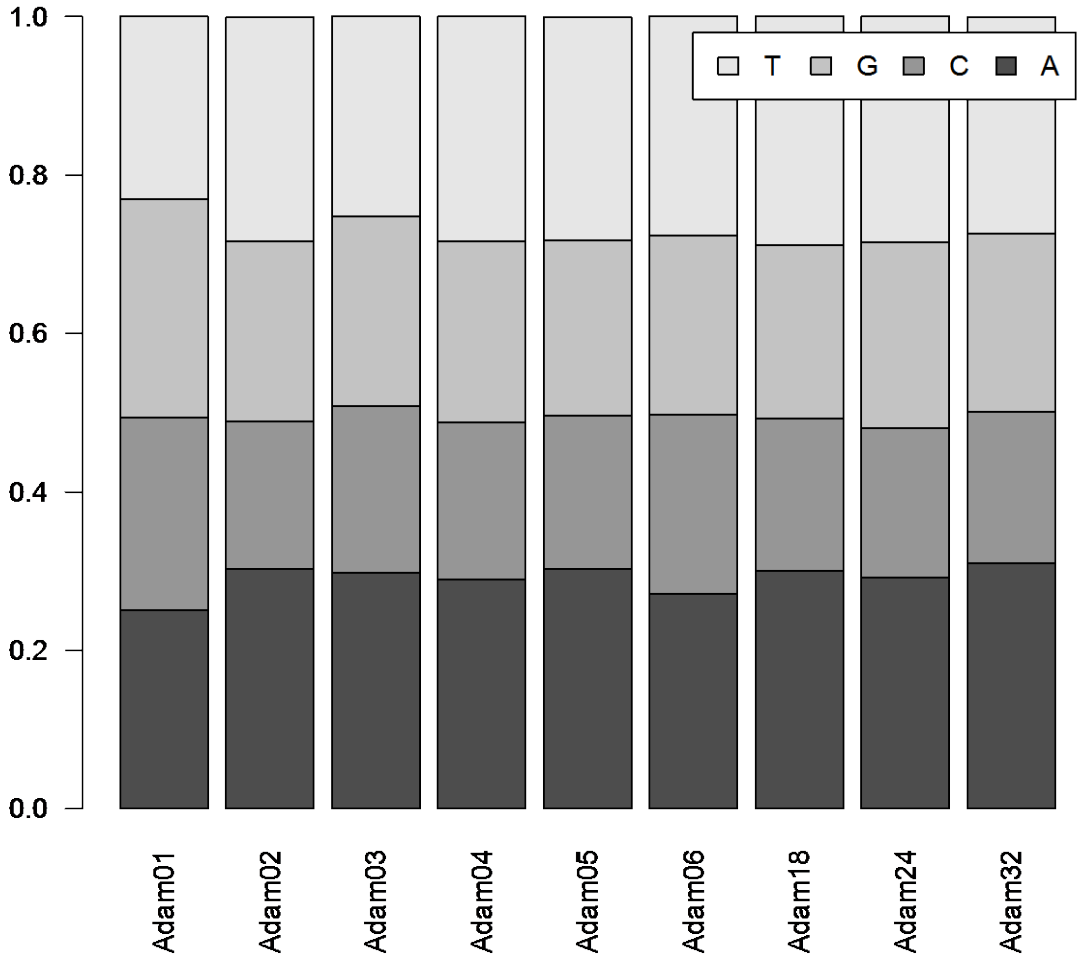

Figure 1. Plot of the base frequencies of the 36 CDSs for the 9 sperm-surface ADAM genes

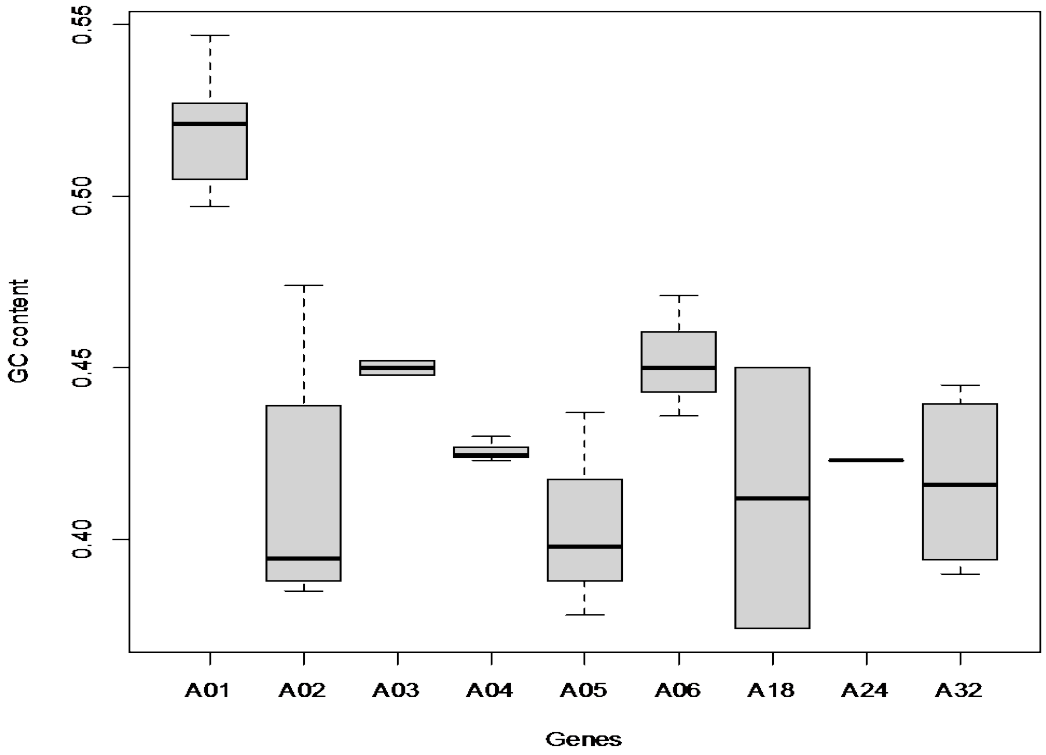

Figure 2. Boxplot of GC Content for the 9 sperm-surface Adam genes. The lines ("whiskers") on the top and bottom of each box show the range of GC content where the horizontal line on each box represents the median

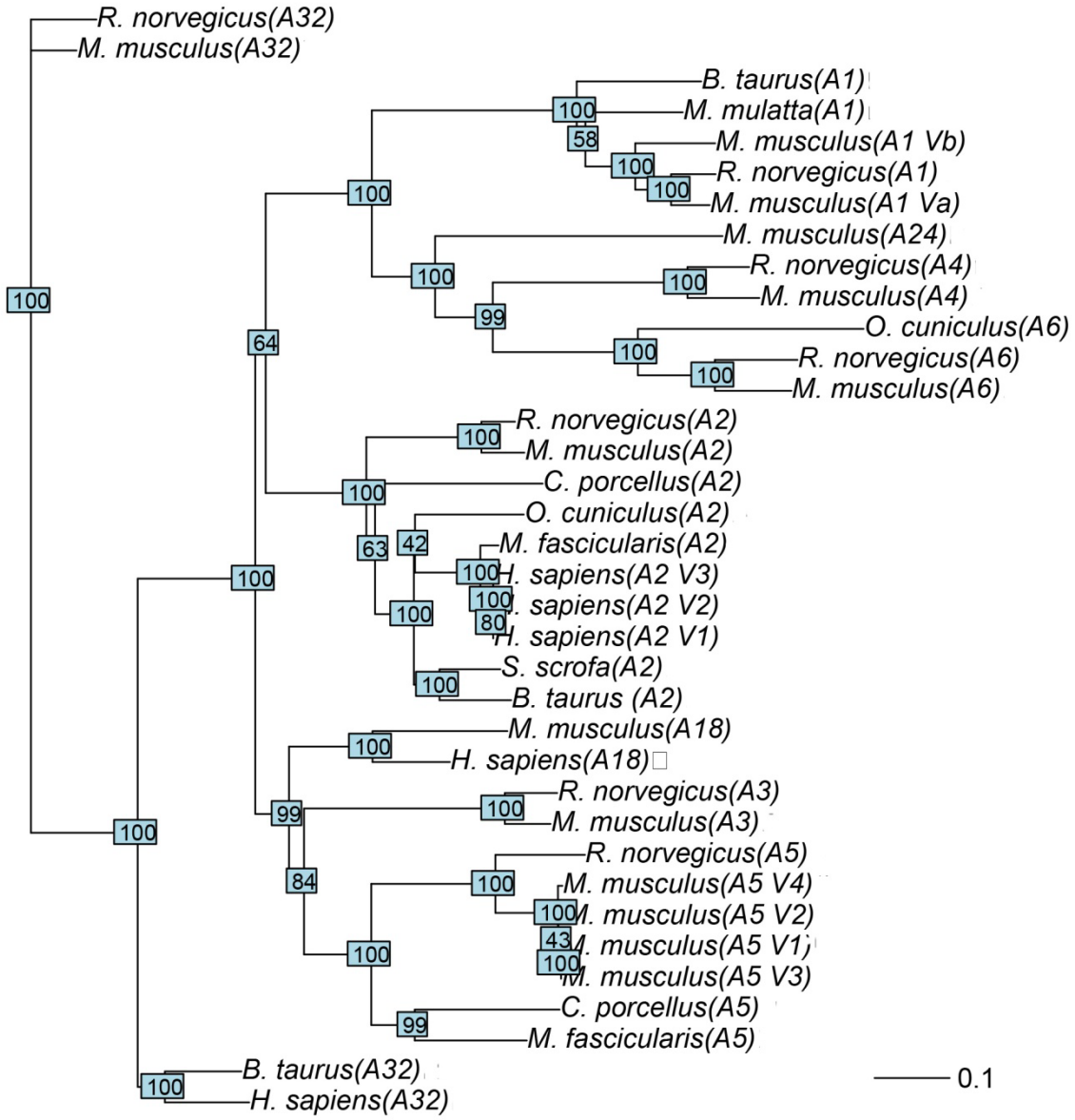

Figure 3. Neighbor-Joining phylogenetic tree of 9 ADAM genes, the scale bar corresponds to 0.1 substitution per site

Table 2. Estimated parameters of polymorphic sites for sperm-surface ADAM genes

	No Sites	No. Mono mrphic sites	No. Poly -morphic sites	Parsimony informative sites	Singleton variable sites
ADAM1	2760	1150	1086	441	645
ADAM2	2313	843	1085	709	376
ADAM3	2472	1918	302	0	302
ADAM4	2301	1882	377	0	377
ADAM5	2371	916	1011	444	567
ADAM6	2271	1250	942	0	942
ADAM18	2223	1596	561	0	561
ADAM32	2427	1267	930	320	610

General information about the polymorphisms on 8 ADAM genes presented in table (2). The number of sites ranged from 2223 for ADAM18 to 2760 for ADAM1, where the smallest number for monomorphic sites was reported for ADAM2 (843) and the largest (1596) was reported for ADAM18. Both ADAM1 & 2 had the largest number of polymorphic sites 1086 & 1085 where ADAM3 had the lowest number of polymorphic sites.

CDS sequences of the 8 ADAM genes were also analyzed to characterize the sequence diversity. The results of the analysis are presented both numerically and graphically (Table 3 & Figure 4). The number of haplotypes was positively related to the number of sequences analyzed per gene, that is, ADAM2 had the largest number of 10 sequences such that it had the largest number of haplotypes, ADAM5 &1 ranked second in number of haplotypes. These three genes were also found to have the largest number of polymorphic sites (table 2). ADAM2, 4, & 18 had the lowest number of sequences and lowest number of haplotypes as well.

Table 3. Estimated parameters of DNA polymorphism for sperm-surface ADAM genes

	No. Haplotypes	Haplotype diversity ±SD	Nucleotide diversity(π)	Average number of nucleotide diff
ADAM1	5	1 ± 0.13	0.27	603.7
ADAM2	8	0.93 ± 0.1	0.23	451.7
ADAM3	2	1 ± 0.5	0.14	302
ADAM4	2	1 ± 0.5	0.21	377
ADAM5	6	0.95 ± 0.1	0.22	432.4
ADAM6	3	1 ± 0.27	0.3	666.3
ADAM18	2	1 ± 0.5	0.26	561
ADAM32	4	1 ± 0.2	0.26	569.6

Estimation of nucleotide diversity (π) showed that not only all analyzed ADAM genes were not equally diverse but also highly variable where π ranged from 0.14 for ADAM3 to 0.3 for ADAM6.

Table (4) shows the conserved regions along the 8 ADAM genes and measurements of conservation (C), homosigoisty and P-value. Conservation (C) is calculated as the proportion of conserved sites in the alignment region, where homosigosity is measured as 1-heterzygosity. All the 8 genes showed significant conserved regions (P >0.05) among the studied sequences with high values of conservation and homozygosity. Similarly to the results in table 3 that wasn't equally diverse, not all genes have equal number of conserved regions neither the length of conserved regions. ADAM1, 2 and 6 have 3 conserved regions but ADAM1 has the longest conserved region (from 633 to 1334) where ADAM18 has the largest number of conserved regions (9). The length of conserved regions ranged from 51 to 701 bp.

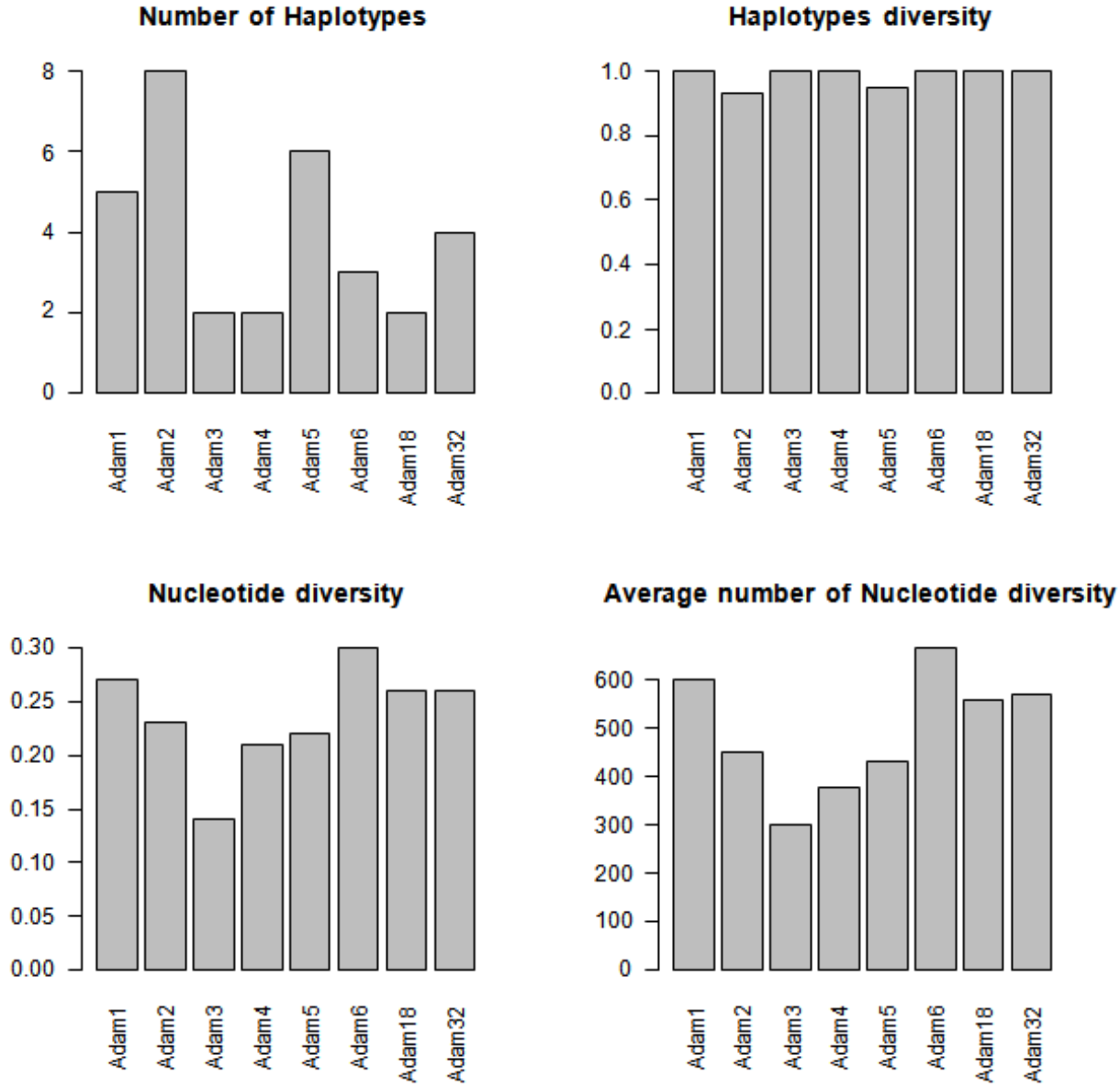

Figure 4. Measures of DNA polymorphism for Adam genes

4. Discussion

To the best of our knowledge this is the first time for sperm-surface ADAM genes to be analyzed based on Refseq data. The present study focused on non-redundant and curated RefSeq data containing 36 CDS of 9 sperm-surface ADAM genes.

Our results show great variation in the length of CDS sequences of sperm- surface ADAM genes either between or within species for example length variation in variants *a* and *b* of ADAM1 in *M. musculus*. Moreover, variation in stop codon were also observed. These results expand on the previous findings on the molecular evolution of ADAM family genes (Glassey & Civetta, 2004) to show that all ADAM genes involved in male reproductive tract show evidence of being under the selection pressure. Dorus et al. (2010) detected positive selection for ADAM1,2,4,6 and 24 using phylogenetic comparisons among five different species of mammals. Moreover, Finn & Civetta (2010) analyzed 25 members of ADAM gene family and found that all genes expressed in male reproductive tissues showed evidence of positive selection. The same study reported positive selection on codon sites within ADAM1, 2, and 32. This signal of positive selection within ADAM genes might be ascribed to species-specific adaptation of fertilization (Shamsadin *et al.*, 1999, Nishimura *et al.*, 2001 and Yamaguchi et al., 2009).

The genomic GC content is one of the key parameters of variation of genome sequences, the value of GC confined to between 25% and 75% (Wu *et al.*, 2012). GC% of the 9 studied ADAM genes ranged between 37 to 55%. This great variation in GC content of sperm-surface ADAM genes is reflection of nucleotide and haplotype diversity (table 3). Choi et al. (2013) reported GC% of 72.9% of non-coding sequences of ADAM2 genes. The length of CDS was shown to be under both functional and structural constrains (Blake, 1983, Blake, 1985, Hawkins, 1988 and Traut, 1988). It is also known that the size distributions of the gene parts (exons, introns, leader and trailer regions, etc.) are under stabilizing selection against extreme lengths (Smith, 1988). Moreover, Oliver & Marin (1996) found that the exon length doesn't affected by concentration of GC in vertebrate. The authors ascribed this result to lower number of exons in the genes located in the studied regions. Our results (table 1) showed that GC% does not affected by the length of CDS for example; the GC% of three variants of ADAM2 of *H. sapeins* did not vary with variation of CDS length. As Oliver & Marin (1996) concluded the effect of GC% on the length of CDS might constitute a new evolutionary meaning for compositional variation in DNA GC content.

Table 4. Length of conserved regions, conservation, homzigosity and P-values of 8 ADAM genes

Gene	Region Start–End	Conservation	Homozigosity	P-value	Gene	Region Start–End	Conservation	Homozigosity	P-value
ADAM1	168–278	0.60	0.77	0.03	ADAM5	1–138	0.57	0.81	0.01
	291–612	0.70	0.83	>0.01		273–445	0.57	0.83	>0.01
	633–1334	0.70	0.84	>0.01		527–743	0.57	0.82	>0.01
						947–1056	0.57	0.83	0.02
ADAM2	236–469	0.57	0.85	>0.01		976–1139	0.58	0.83	>0.01
	563–1112	0.54	0.83	>0.01		1171–1266	0.58	0.83	0.02
	2142–2295	0.54	0.83	0.002					
					ADAM6	1–543	0.71	0.80	>0.01
ADAM3	151–214	0.95	0.97	0.01		1346–1637	0.68	0.78	>0.01
	667–726	0.98	0.98	0.002		1688–1884	0.69	0.79	>0.01
	820–899	0.98	0.98	0.001					
	936–987	0.98	0.98	0.004	ADAM18	161–255	0.83	0.83	0.02
	1528–1632	0.97	0.97	>0.01		182–260	0.84	0.84	0.03
	1864–1915	0.98	0.98	>0.01		187–267	0.84	0.84	0.02
						271–435	0.84	0.84	>0.01
ADAM4	107–308	0.95	0.95	>0.01		875–950	0.84	0.84	0.02
	331–568	0.95	0.95	>0.01		1397–1478	0.84	0.84	0.02
	963–1028	0.96	0.96	0.003		1863–1937	0.84	0.84	0.03
	1096–1150	0.95	0.95	0.01		1864–1943	0.84	0.84	0.03
	1447–1500	0.94	0.94	0.01		1885–1979	0.83	0.83	0.02
	1556–1676	0.94	0.93	>0.01					
	1813–1900	0.93	0.93	0.01	ADAM32	26-471	0.69	0.81	>0.01
						560–1077	0.69	0.83	>0.01
						1096–1192	0.68	0.82	0.02
						1527–1610	0.68	0.78	0.03
						1544–1647	0.67	0.79	0.02
						1565–1664	0.68	0.79	0.02

Although ADAM genes are conserved in evolution, not all members of this family are not found in some mammals (Long *et al.*, 2012). In the present study we investigated the phylogenetic of sperm-surface ADAM genes which is important of in understanding the evolutionary history of these genes. We found that the phylogeny of sperm-surface genes was well supported. Our phylogenetic reconstruction supported the overall orthology of sperm-surface ADAM genes. ADAM1,4,6 and 24 were grouped together in one clade. This clade did not include *H. sapiens* which indicated these genes do not have human orthologues. This result is in agreement with the finding of Choi et al. (2004). However, the only differences in the topology of tree between our results and the outputs of both Finn & Civetta (2010) and Grayson & Civetta (2013) is inclusion of ADAM32 of all species in the same clade. Finn & Civetta (2010) used redundant data from several species where Finn & Civetta (2010) worked only on Mus species. In fact these results shade the light into the significance of examining different factors such as selection within specific groups or clades, because the effect of selection might be impaired by phylogenetic analysis that include various species. The same conclusion was drawn by Civetta (2012) and Garyson & Civetta (2013).

Our results extend to previous finding that considerable sequence diver- sity exists among sperm-surface ADAM genes. We found that this was a reflected by the nucleotide diversity and the average number of nucleotide differences. Our analysis revealed that individual members of sperm-surface ADAM genes differs widely in their average number of haplotypes, nucleotide diversity and the average number of nucleotide differences. Moreover, not all polymorphic sites were informative where, ADAM3,4,6 and 18 were lacking informative sites.

The analysis of the conserved regions of sperm-surface ADAM genes sup-ported the orthology within each member of this gene family. This conclusion is justifiable with the high estimates of conservation, homozigosity and P-values. The conservation within each of ADAM genes is in a good agreement with their essential functionality *in vivo* as determined by knocked out mice. In conclusion, the *in silico* analysis of the 9 sperm-surface ADAM genes showed a great deal of variation among and within this genes indicating the presence of localized signals of selection pressure on these genes. Moreover, ADAM1,4,6 and 24 have no human orthologues. More importantly, this result suggests that ADAM1,2,3,4,5,6, 18,24 and ADAM32 of *B.taurus* and *H. sapiens* were descendent form one ancient ancestor, where ADAM32 of *R. norvegicus* and *M. musculus* have another ancestor.

REFERENCES

[1] A, Civetta. 2003. Shall we dance or shall we fight? Using DNA sequence data to untangle controversies surrounding sexual selection. Genome, 46(925–929).

[2] Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., & Wheeler, D.L. 2005. GenBAnk. Nucleic Acids Res, 33, D34–D38.

[3] Black, R.A., & White, J.M. 1998. ADAMs: focus on the protease domain. Curr Opin. Cell Biol, 10, 654–659.

[4] Blake, C. 1983. Exones – present form the beginning? Nature, 306, 535–537.

[5] Blake, C. 1985. Exons and the evolution of proteins. Int. Rev Cytol, 93, 149–185.

[6] Blobel, C.P. 1997. Metalloprotease-disintegrins: links to cell adhesion and cleavage of TNF alpha and Notch. Cell, 90, 589–592.

[7] Blobel, C.P., Wolfsberg, T.G., Turck, C.W., Myles, D.G., P., Primakoff., & White, J.M. 1992. A potential fusion peptide and an integrin ligand domain in a protein active in sperm-egg fusion. Nature, 356, 284–252.

[8] Cho, C. 2005. Mammalian ADAMS with Testis-specific or Predominant Expression. Chap. 11, pages 239–258 of: Hooper, N. M., & Lendeckel, U. (eds), The ADAM Family of Proteases. The Netherlands: Springer.

[9] Cho, C. 2012. Testicular and epididymal ADAMs: expression and function during fertilization. Nat. Rev. Urol., 9, 550–560.

[10] Cho, C., Bunch, D.O., Faure, J.E., Goulding, E.H., Eddy, E.M., Primakoff, P., & Myles, D.G. 1998. Fertilization defects in sperm from mice lacking fertilin beta. Science, 281, 1857–1879.

[11] Choi, H., Lee, B., Jin, S., Kwon, J. T., Kim, J., Oh, S., Cho, B-N., Prk, Z. Y., & Cho, C. 2013. Idenificatin and characherization of promoter and regulatory regions for muse Adam2 gene expression. Mol. Biol. Rep., 40, 787–796.

[12] Choi, I., Oh, J aand Cho, B-N, Ahnn, J., Jung, Y-K, Kim, D. H., & Cho, C. 2004. Characterization and comparative genomic analysis of intronless Adams with testicular gene expression. Genomics, 83, 636–646.

[13] Civetta, A. 2003. Positive selection within sperm-egg adhesion domains of fertilin: an ADAM gene with a potential role in fertilization. Mol Biol Evol, 20, 21–29.

[14] Civetta, A. 2012. Fast evolution of reproductive genes: when is selection sexual? Pages 165–175 of: Singh, R., Xu, J., & Kulathinal, R. (eds), Rapidly evolving genes and genetic systems. Oxford, UK: Oxford University Press.

[15] Clark, N.L., Aagaard, J.E., & Swanson, W.J. 2006. Evolution of reproductive proteins from animals and plants. Reproduction, 131, 11–22.

[16] Dorus, S., Wasbrough, E.R., Busby, J., Wilkin, E.C., & Karr, T.L. 2010. Sperm proteomics reveals intensified selection on mouse sperm membrane and acrosome genes. Mol Biol Evol, 27, 1235–1246.

[17] Evans, J.P. 2002. The molecular basis of spermoocyte membrane interactions during mammalian fertilization. Hum Reprod Update, 8, 297–311.

[18] Finn, S., & Civetta, A. 2010. Sexual Selection and the Molecular Evolution of ADAM Proteins. J. Mol. Evol., 71, 231–240.

[19] Fisher, H. S., & Hoekstra, H.E. 2010. Competition drives cooperation among closely related sperm of deer mice. Nature, 463, 801–803.

[20] Frayne, J., Hurd, E.A.C., & Hall, L. 2002. Human tMDC III: a sperm protein with a potential role in oocyte recognition. Mol Hum Reprod, 8, 817–822.

[21] Garyson, P., & Civetta, A. 2013. Positive selection in the adhesion domain of Mus sperm Adam genes through gene duplications and function- driven gene complex formations. BMC Evolutionry Biology, 13, 217–224.

[22] Glassey, B., & Civetta, A. 2004. Positive selection at reproductive ADAM genes with potential intercellular binding activity. Mol Biol Evol, 21, 851–859.

[23] Grayson, P., & Civetta, A. 2013. Positive selection in the adhesion domain of Mus sperm Adem gene duplications adn fuction driven gene complex formations. BMC Evolutionary Biology, 13, 217–224.

[24] Han, C., Kwon, J.T., Park, I., Lee, B., Jin, S., Choi, H., & Cho, C. 2010. Impaired sperm aggregation in Adam2 and Adam3 null mice. Fertil Steril, 93, 2754–2756.

[25] Hawkins, J.D. 1988. A survey on intron and exon lengths. Nucleic Acids Res., 16(9893–9908).

[26] Kim, E., Yamashita, M., Nakanishi, T., Park, K-E., Kimura, M., Kashiwabara, S., & Baba, T. 2006. Mouse sperm lacking ADAM1b/ADAM2 fertilin can fuse with the egg plasma membrane and effect fertilization. J Biol Chem, 261, 5634–5639.

[27] Kim, T., Oh, J., Woo, J.M., Choi, E., Im, S.H., Yoo, Y.J., Kim, D.H., Nishimura, H., & Cho, C. 2005. Expression and relationship of male reproductive ADAMs in mouse. Biol Reprod, 74, 744–750.

[28] Koonin, E.V., & Wolf, Y.I. 2010. Constraints and plasticity in genome and molecular- phenome evolution. Nat Rev Genet, 11, 487–489.

[29] Kosiol, C., Vinar, T., Da-Fonseca, R.R., Hubisz, M.J., Bustamante, C.D., & Siepel, A. 2008. Patterns of positive selection in six mammalian genomes. PLoS Genet, 4, e1000144.

[30] Long, J., Meng Li, M., Ren, Q., Zhang, C., Fan, J., Duan, Y., Chen, J., Li, B., & Deng, L. 2012. Phylogenetic and molecular evolution of the ADAM (A Disintegrin And Metalloprotease) gene family from Xenopus tropicalis, to Mus musculus, Rattus norvegicus, and Homo sapi- ens. Gene, 507, 36–43.

[31] Lynch, M., & Crease, T.J. 1990. The analysis of population survey data on DNA sequence variation. Mol Biol Evol, 7, 377–394.

[32] Moore, H., Dvorakova, K., Jenkins, N., & Breed, W. 2002. Excep- tional sperm cooperation in the wood mouse. Nature, 418, 174–177.

[33] Morgan, C.C., Loughran, N.B., Walsh, T., Harrison, A.J., & OConnell, M.J. 2010. Positive selection neighboring functionally essen- tial sites and disease-implicated regions of mammalian reproductive proteins. BMC Evolionary Biology, 10, 39–57.

[34] Nishimura, H., Cho, C., Branciforte, D.R., Myles, D.G., & Primakoff, P. 2001. Analysis of loss of adhesive function in sperm lacking cyritestin or fertilin beta. Dev. Biol., 233, 204–213.

[35] Nishimura, H., Kim, E., Nakanishi, T., & Baba, T. 2004. Possible function of the ADAM1a/ ADAM2 Fertilin complex in the appearance of ADAM3 on the sperm surface. J Biol Chem, 279, 34957–34962.

[36] Oliver, J. L., & Marin, A. 1996. A relationship between GC content and coding-Sequence length. Journal of Molecular Evolution, 43, 216–223.

[37] Pages, H., Aboyoun, P., Gentleman, R., & DebRoy, S. 2013. Biostrings: String objects representing biological sequences, and matching algorithms. R package version 2.30.1.

[38] Panhuis, T.M., Clark, N.L., & Swanson, W.J. 2006. Rapid evolution of reproductive proteins in abalone and Drosophila. Philos T Roy Soc B, 361, 261268.

[39] Paradis, E., Claude, J., & Strimmer, K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics, 20, 289–290.

[40] Primakoff, P., & Myles, D.G. 2000. The ADAM gene family: surface proteins with adhesion and protease activity. Trends Genet, 16, 83–87.

[41] Pruitt, K.M, Tatusova, T., & Maglott, D.R. 2005. NCBI Refer- ence Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Research, Database issue 33, D501–D504.

[42] R Core Team. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

[43] Saitou, N, & Nei, M. 1987. The neighbor-joining method: a new method for reconsructing phylgentic trees. Mol. Biol. Evol., 4, 406–425.

[44] Seals, D.F., & Courtneidge, S.A. 2003. The ADAMs family of metalloproteases: multidomain proteins with multiple functions. Genes Dev, 17, 7–30.

[45] Shamsadin, R., Adham, I.M., Nayernia, K., Heinlein, U.A., Oberwinkler, H., & Engel, W. 1999. Male mice defi t for germ-cell cyritestin are infertile. Biol Reprod, 61, 1445–1451.

[46] Smith, W. M. 1988. Stuctural of vertebrate genes: a statistical analysis iplication selection. J. Mol. Evol., 27, 45–55.

[47] Swanson, W.J., & Vacquier, V.D. 2002. The rapid evolution of reproductive proteins. Genetics, 3, 137–144.

[48] Tajima, F. 1983. Evolutionary relationship of DNA sequences in fi populations. Genetics, 105, 437–460.

[49] Traut, T. W. 1988. Do exons code for structural or functional units in proteins? Proc. Natl Acad. Sci. USA, 85, 2944–2948.

[50] Turner, L.M., & Hoekstra, H.E. 2008. Causes and consequences of the evolution of reproductive proteins. Int J Dev Biol, 52, 769780.

[51] Voight, B.F., Kudaravalli, S., Wen, X., & Pritchard, J.K. 2006. A map of recent positive selection in the human genome. PLoS Biol, 4, 446–458.

[52] Wolfsberg, T. G., Bazan, J. J., Blobel, C. P., Myles, D. G., Primakoff, P., & White, J. M. 1993. The precursor region of a protein active in sperm-egg fusion contains a metalloprotease

and a disintegrin domain: Structural, functional and evolutionary implications. Proc. Natl. Acad. Sci. USA, 90, 10783–10787.

[53] Wolfsberg, T.G., Primakoff, P., Myles, D.G., & White, J.M. 1995a. ADAM, a novel family of membrane proteins containing a disintegrin and metalloprotease domain: multipotential functions in cellcell and cellmatrix interactions. J Cell Biol, 131, 275–278.

[54] Wolfsberg, T.G., Straight, P.D., Gerena, R.L., Huovila, A.P., Primakoff, P., Myles, D.G., & White, J.M. 1995b. ADAM, a widely distributed and developmentally regulated gene family encoding membrane proteins with a disintegrin and metalloprotease domain. Dev. Biol, 169, 378–383.

[55] Wu, H., Zhang, Z., Hu, S., & Yu, J. 2012. On the molecular mechanism of GC content variation among eubacterial genomes. Biology Direct, 7, 2– 15.

[56] Yamaguchi, R., Muro, Y., Isotani, A., Tokuhiro, K., Takumi, K., Adham, I., Ikawa, M., & Okabe, M. 2009. Disruption of ADAM3 impairs the migration of sperm into oviduct in mouse. Biol Reprod, 81, 142–146.

[57] Zhu, G.Z., S., Gupta., Myles, D.G., & Primakoff, P. 2009. Testase 1 (ADAM 24) a sperm surface metalloprotease is required for normal fertility in mice. Mol Reprod Dev, 76(1106–1114).